Genes, variant genes, and pseudogenes of the human tRNA^{Val} gene family are differentially expressed in HeLa cells and in human placenta

Cornelia Schmutzler[+] and Hans J.Gross*

Institut für Biochemie, Bayerische Julius-Maximilians-Universität, Röntgenring 11, D-8700 Würzburg, FRG

ABSTRACT

Pre-tRNAs^{Val} were identified in unfractionated tRNA preparations from HeLa cells and human placenta and their 5' leader structures were deduced from the nucleotide sequences of the corresponding cDNAs. Several of these precursors can be assigned to nine out of the eleven members of the human tRNA^{Val} gene family characterized so far, which demonstrates that these gene loci are actively transcribed in vivo. Among the expressed genes there are (a) genes for the two known tRNA^{Val} isoacceptor species from human placenta, (b) gene variants that exhibit sequence alterations as compared to conventional genes, and (c) pseudogenes that produce processing-deficient precursors which are not matured to tRNAs. The transcription products of several yet unknown tRNA^{Val} genes have also been detected. Furthermore, different expression patterns are observed in the two cell types studied. These data allow for the first time an insight into the in vivo expression of a human tRNA gene family.

INTRODUCTION

Recently, we reported the cloning and sequencing of nine members of the human tRNA^{Val} gene family and summarized some of their functional properties in vitro (1−3). Pirtle and coworkers characterized two human gene clusters, λ hPTV5 (4) and λ hVKV7 (5), which contain, apart from genes for other tRNA species, one and two tRNA^{Val} genes, respectively. One of the two tRNA^{Val} genes in the latter clone seems to be identical with a tRNA^{Val} gene also described by us except for some point mutations in its 5' flanking region. Thus, there is information available about altogether eleven human tRNA^{Val} genes. This is at least one half of their total number, since human tRNA multigene families are estimated to contain about 20 members each (6).

The composition of this gene family, as elucidated so far, provides some surprise, because considerable sequence variability

is detected among its members. Three genes, those in clones pHtV1 (1), pHtV6, and pHtV8 (3), the latter of which corresponds to the tRNA$^{Val}_{AAC}$ gene in clone λ hVKV7 (5), encode the major tRNA$^{Val}_{IAC}$, and two genes, one in clone pHtV3 (1) and one in clone λ hVKV7 (5), encode the minor tRNA$^{Val}_{CAC}$ from human placenta. Hence, only five of the eleven human tRNA^{Val} genes characterized so far code for one of the known functional valine-specific isoacceptors (7). The other six genes exhibit sequence differences in the region coding for the mature tRNA.

In three cases, i. e., the tRNA^{Val} genes in clones pHtV5, pHtV7 (3) and λ hPTV5 (4), the mutations do not seem to impair the structure and function of the corresponding tRNA products, since conserved, invariable nucleotides are not concerned and since the formation of a normal two- and threedimensional tRNA shape still appears to be possible. In vitro, these three genes are transcribed and processed to give products of the size and the sequence expected for a mature tRNA (3,4). The name 'allogenes' has been proposed for such genes which may be functional in vivo and code for yet uncharacterized minor isoacceptors (8). It should be mentioned that a third valine-accepting tRNA species is known to exist in human placenta which has not been purified and sequenced so far (9).

The three remaining genes have to be regarded as pseudogenes. Although they are functional in the sense that they are transcribed in vitro, they do not lead to the formation of mature tRNAs. In two cases, i. e., the genes in clones pHtV4 and pHtV9 (2,3), the respective mutations presumably cause crippled secondary and tertiary structures of the corresponding transcripts. Consequently, these are not processed at all, because maturation enzymes, which obviously depend on a correct tertiary structure of the mature tRNA domain (10), do not accept them as substrates. As for the third pseudogene, the one in clone pHtV2, the mutations do not affect the secondary or tertiary structure of its RNA products. The primary transcript undergoes a certain degree of maturation and the 5' end is correctly cleaved, however, the 3' end is not correctly processed and the CCA terminus can not be added (3).

This remarkable structural and functional diversity within a human tRNA gene family, showing redundancy as well as variety of information, prompts the question what may be its significance for the organism. Moreover, the knowledge of eleven tRNA^{Val} genes is a unique prerequisite for studying the expression of this tRNA gene family *in vivo*. Interesting issues in this context may concern:

a) the expression and regulation of the gene family in general or of its single representatives in particular and how they contribute to maintain or alter the pool of mature tRNAs in different cells, tissues or differentiation stages;

b) the role of the variant genes also with respect to the question of developmental stage or tissue specificity, as there is already evidence for the occurrence of variant tRNAs in highly specialized organs as, e. g., the silkgland of *Bombyx mori*; in the case of the most extensively studied example, tRNA^{Ala}, this is due to the activation of a tissue-specific tRNA gene (11);

c) the activity and possible functions of pseudogenes.

Unfortunately, such studies are impaired by the already mentioned redundancy of genetic information for a certain isoacceptor species, a characteristic property of eukaryotic tRNA gene families. Within the human tRNA^{Val} gene family this redundancy has been demonstrated for the major tRNA^{Val}$_{IAC}$ as well as for the minor tRNA^{Val}$_{CAC}$ with at least three and two genes coding for them, respectively. Hence, the presence of a certain tRNA in cells or tissues provides no evidence for the expression of one distinct gene out of a group of genes coding for the same isoacceptor. The activity of a single gene locus can be demonstrated if, e. g., tRNA precursors deriving from it are detected. As the flanking regions of tRNA genes mostly differ considerably (for the tRNA^{Val} genes compare refs. 3—5), the knowledge of the 5' leader sequence of a precursor should suffice to assign the latter to a single tRNA gene locus. However, tRNA precursors are certainly subject to a rapid turnover by the processing enzymes and presumably occur in only very low concentrations. Consequently, any experimental technique used must be able to detect very small amounts of pre-tRNAs in the presence of a high excess of mature tRNA.

We found that primer extension meets this necessity. By applying this technique we could demonstrate the activity of several single tRNA^{Val} gene loci in HeLa cells and human placenta. Among these, there are genes, variant genes and pseudogenes which are already well characterized (1—5) as well as yet unknown tRNA^{Val} genes. Thus, this allowed for the first time to get an insight into the expression of a human tRNA gene family *in vivo*.

MATERIALS AND METHODS

Enzymes and Chemicals

T4 polynucleotide kinase was purchased from NEN. AMV reverse transcriptase, salmon sperm DNA, and dideoxynucleoside triphosphates were from Boehringer, Mannheim. (γ-^{32}P)-ATP with a specific activity of 110 Tbq/mmol was obtained from Amersham. Guanidinium rhodanide was from Fluka and DEAE cellulose from Whatman. All other chemicals were obtained from Fluka, Merck, Serva, or Sigma. Oligodeoxyribonucleotide primers for reverse transcription were synthesized using the DNA synthesizer 2000 from Applied Biosystems. Pure tRNA^{Val}$_{IAC}$ from rabbit liver was prepared in this laboratory (12).

Isolation of low molecular weight RNA

RNA was isolated by lysis of a HeLa cell pellet or of placenta tissue in the presence of guanidinium rhodanide according to

Maniatis *et al.* (13). The high molecular weight fraction was removed from the RNA preparations by DEAE cellulose chromatography essentially as described by Roe (14). The low molecular weight fraction contains mainly 5S rRNA and tRNA.

5'-labelling of primer oligonucleotides

25 pmoles of the primer, 500 μCi of (γ-^{32}P)-ATP and 5 units of T4 polynucleotide kinase were incubated for 30 min at 37°C in 10 μl reaction buffer (100 mM Tris·HCl, pH 8.0; 20 mM MgCl$_2$; 3.3 mM spermin; 33 mM β-mercaptoethanol). The 15-mer or 20-mer oligodeoxyribonucleotides (Fig. 1) were separated from shorter oligomers and from unincorporated radioactivity on 20% polyacrylamide/8 M urea gels.

Primer extension

Primer extension reactions on a preparative scale were done as follows: 60 to 150 μg low molecular weight RNA from Hela cells or from human placenta and 3×10^7 to 10^8 cpm (10 to 30 ng; 2 to 6 pmoles) 5' labelled primer oligonucleotide were precipitated with 0.1 vol 2 M sodium acetate, pH 5.2 and 2 vol of ethanol. For primer hybridization, the precipitated nucleic acids were dissolved in 100 μl buffer I (10 mM Tris·HCl, pH 8.3; 250 mM NaCl; 1 mM EDTA), heated to 95°C for 10 min, cooled on ice for 5 min and incubated at room temperature for 20 min.
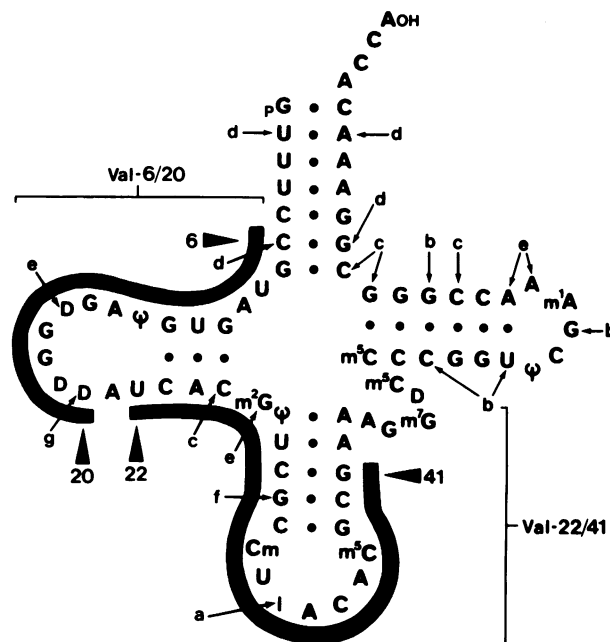
Figure 1: Oligonucleotide primers for reverse transcription of tRNAs^{Val} and pre-tRNAs^{Val}. The primer oligonucleotides are represented by black bars. Oligonucleotide Val-6/20 (nucleotide sequence $^{5'}$AACCACTACACTACG$^{3'}$), spanning positions 6 to 20 of tRNA^{Val}, was used for most tRNA^{Val}- and pre-tRNA^{Val}-primed cDNA syntheses. Primer Val-22/41 (nucleotide sequence $^{5'}$CGCGTGTTAGGCGAACGTGA$^{3'}$), homologous to positions 22 to 41 of tRNA^{Val}, was used for the detection of pre-tRNAs deriving from genes which carry mutations in or beyond the sequences hybridizing to primer Val-6/20. The sequence of the mature major tRNA^{Val}$_{IAC}$ species is shown. The cellular tRNA^{Val} genes, gene variants and pseudogenes represented by the following clones carry mutations of the region coding for the mature tRNA domain in the positions indicated by small letters. (a): pHtV3, pHtV4, pHtV5, pHtV7, pHtV9, and the tRNA^{Val}$_{CAC}$ gene in clone λ hVKV7 carry a C in this position. (b): pHtV2 has a U$_{51}$, an A$_{54}$, an A$_{57}$, and an A$_{63}$. (c): pHtV4 carries a U$_{25}$, a U$_{62}$, an A$_{65}$, and a U$_{60}$. (d): pHtV5 has a C$_2$, a U$_6$, an A$_{67}$ and a G$_{71}$. (e): pHtV7 has a C$_{16}$, an A$_{26}$, a U$_{59}$, and a C$_{60}$. (f): pHtV9 has an A$_{30}$. (g): the variant tRNA^{Val}$_{AAC}$ gene in λ hPTV5 has a C$_{20}$.

After addition of 150 μl buffer II (25 mM Tris·HCl, pH 8.3; 16 mM MgCl₂; 0,4 mM each dATP, dCTP, and TTP; 0,8 mM dGTP; 8 mM DTT) and 30 units of reverse transcriptase, primer extension was performed for 60 min at 37°C.

To stop the elongation reaction, the samples were heated to 95°C for 5 min. After ethanol preciptation, the nucleic acids were dissolved in 50 μl loading buffer (90% formamide, 10 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol), denatured for 10 min at 95°C, and cooled on ice for 2 min. The short elongation products obtained by priming with oligonucleotide Val-6/20 (Fig. 1) were purified on 25% polyacrylamide/6 M urea gels (dimensions: 880×240×0.35 mm; slot width: 4 cm; 45 h at 3000 V and 20 mA). Longer reverse transcripts synthesized after priming with Val-22/41 (Fig. 1) were separated on 20% polyacrylamide/8 M urea gels (dimensions: 600×200×0.35 mm; slot width: 4 cm; 10 h at 2500 V and 17 mA). Following electrophoresis, the gels were exposed for different time intervals on preflashed Fuji RX X-ray film in the presence of an intensifying screen.

Using the autoradiograms, reverse transcription products of interest were localized, excised from the gel, and eluted by diffusion into 200 μl of 10 mM Tris·HCl, 1 mM EDTA, pH 7.5. They were precipitated with ethanol in the presence of 10 μg salmon sperm DNA, redissolved in 32 μl of sterile H₂O and used for sequencing by the method of Maxam and Gilbert (15).

Control reactions on an analytical scale contained 1 μg purified rabbit liver tRNA^{Val}, poly(A) as a carrier for the precipitation (20 μg), 3 U of reverse transcriptase, and 10⁶ cpm of labelled primer oligonucleotide in 10 μl of buffer I and 15 μl of buffer II (see above).

DNA sequencing

Chemical sequencing reactions were performed essentially as described by Maxam and Gilbert (15). However, incubation times were varied in the following manner to achieve sufficient cleavage of the short DNA molecules analyzed: G-reaction, 10 min/20°C; A+G-reaction, 20 min/20°C; C+T-reaction and C-reaction, 1 h/37°C. Sequencing gels contained 15% or 20% polyacrylamide and 8 M urea.

RESULTS

### Design of suitable primers for primer extension with pre-tRNAs^{Val} as templates

For most primer extension experiments we used the pentadecadeoxyribonucleotide Val-6/20 (Fig. 1) which hybridizes to nucleotides 6 to 20 of the mature tRNAs^{Val}. Elongation of this oligonucleotide with tRNA^{Val} as a template causes an extension of five nucleotides which are complementary to the first five nucleotides of the mature tRNA. In case of a tRNA^{Val} precursor, the elongation product is, in addition, complementary to the nucleotides of the 5′ leader. Transcription initiation of all tRNA^{Val} genes examined in our laboratory occurs between positions −2 and −9, at least *in vitro* (3). This leads to primer extension products 20 to 30 nucleotides long which can be separated from each other by electrophoresis on long 25% polyacrylamide/6 M urea gels. A primer length of fifteen nucleotides ensures sufficient specifity of hybridization with the tRNA^{Val} or pre-tRNA^{Val} templates; any RNA or gene sequence examined using the NIH GenBank data showed less than 10 nucleotides homology with this oligonucleotide. Hence, only inefficient hybridization to other RNAs that could impair the

specificity of the primer extension reaction is expected. The fact that the region of complementarity between the oligonucleotide and tRNA^{Val} starts at nucleotide 6 of the mature tRNA and not at nucleotide 1 was intended as an additional control for the specificity of the hybridization. In mammalian tRNAs (or tRNA genes), the sequence motif 5′GUUUC...3′ (5′...GTTTC...3′) has so far exclusively been found at the 5′ end of tRNAs^{Val} (at the 5′ end of genes coding for tRNAs^{Val})(16). Therefore, if the sequence 5′...GAAAC...3′ is detected in a reverse transcript, this should serve as sufficient evidence that the corresponding template was a pre-tRNA^{Val}.

Some experiments were done in order to visualize mutations in the DHU-arm of variant tRNA^{Val} species. The primer used in this case was the ikosadeoxyribonucleotide Val-22/41 which hybridizes to the anticodon stem and loop and part of the DHU-stem (Fig. 1).

Preparative primer extension

Most of the primer extension experiments performed in this study were done with the intention to synthesize the cDNAs of pre-
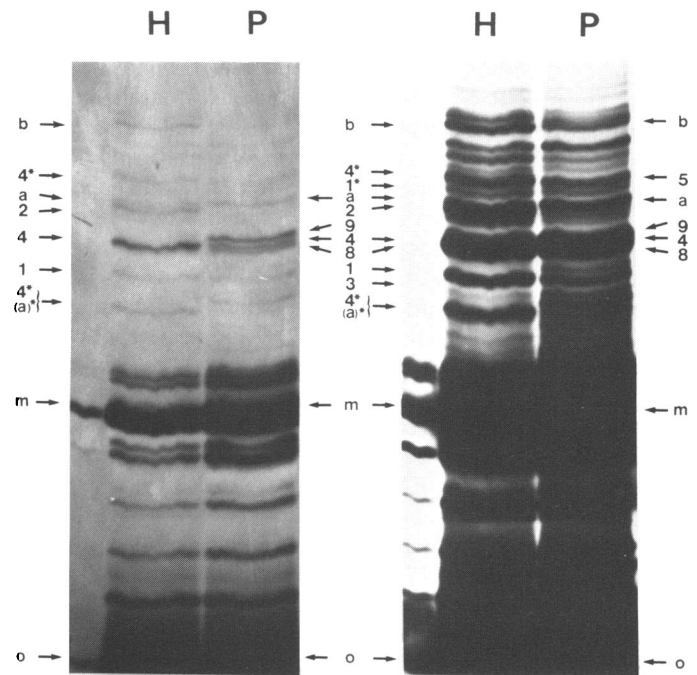
Figure 2: Reverse transcripts of pre-tRNAs^{Val} from HeLa cells and human placenta. Low molecular weight RNA from HeLa cells (H) and placenta (P) were reverse transcribed using the 5′-(³²P)-labelled primer Val-6/20, and the cDNAs were separated on a 88 cm long preparative 25% polyacrylamide/8 M urea gel. Autoradiography was for 2 h (left panel) and 15 h (right panel), respectively. The narrow lane left of the HeLa cDNAs shows the mobility of the primer oligonucleotide used (o) and of the elongation product obtained in a control experiment with purified, mature tRNA^{Val} (m). Numbers and arrows identify cDNAs deriving from pre-tRNAs of respective cellular tRNA^{Val} genes corresponding to clones pHtV1 (= 1) to pHtV9 (= 9). cDNAs of pre-tRNAs transcribed from pHtV6- and pHtV7-related genes, respectively, were not detected. The cDNAs of pre-tRNAs of the gene related to the tRNA$^{Val}_{CAC}$ gene in clone λ hVKV7 (5) and of a yet unknown tRNA^{Val} gene are labelled 'a' and 'b', respectively. '*' marks additional, minor initiation sites, and the brace signifies that the pre-tRNA which gave rise to this cDNA may have been transcribed either from the gene corresponding to pHtV4 or from that related to the tRNA$^{Val}_{CAC}$ gene in clone λ hVKV7 (initiation at −3, compare Table 1). Expression of the tRNA^{Val} gene corresponding to that in clone λ hPTV5 was detected in an experiment with the primer Val-22/41 (not shown). Note that bands 4* (H) and 5 (P) in the right panel are not identical in spite of their almost identical electrophoretic mobility.

tRNAs^{Val} in sufficient amounts to determine their nucleotide sequences according to the method of Maxam and Gilbert (15). For this purpose, low molecular weight RNA from HeLa cells or from human placenta was subjected to primer extension (preparative scale) with the oligonucleotide Val-6/20. In parallel, a control reaction (analytical scale) with purified tRNA^{Val} from rabbit liver was carried out. Conditions for the assay and the electrophoretic separation of the elongation products are described in Materials and Methods. Fig. 2 shows an autoradiogram produced in the course of such an experiment. In the control reaction the longest reverse transcription product is the one which is obtained upon primer elongation with mature tRNA^{Val} as a template (Fig. 2, m). In the assays with unfractionated tRNA preparations (Fig. 2, lanes H and P) there are additional, longer cDNAs with lower electrophoretic mobility which derive from tRNA precursor molecules. The separation achieved by the 88 cm long 25% polyacrylamide/6 M urea gel is not only according to the length, but also partly according to the nucleotide composition of the primer extension products. Therefore, a considerable number of precursor-derived products is dissolved in lanes H and P. The number of bands which are clearly distinguishable from each other are about thirty in each of the two RNA preparations.

Primer extension products deriving from precursors as well as from mature tRNA^{Val} were excised, eluted from the gel, and the radioactivity was determined in a liquid scintillation counter. If one assumes that the yield of reverse transcript is directly proportional to the amount of the corresponding template, the ratio of the mature tRNA to pre-tRNAs in an RNA preparation analyzed by primer extension can be calculated from the respective ^{32}P counts. Usually about 3.5×10^7 and 5×10^7 cpm were incorporated into the cDNAs deriving from mature tRNA in the case of HeLa cell and placenta RNA, respectively, whereas the more or less prominent products derived from precursor molecules ranged from several thousand to about 6×10^4 cpm. Therefore we conclude that in HeLa cells or placenta tissue the mature tRNA^{Val} is present in an at least 500-fold excess over the corresponding pre-tRNAs^{Val}.

Eluted cDNAs containing more than 4000 cpm were subjected to nucleotide sequence analysis by the chemical cleavage method of Maxam and Gilbert (15). In preliminary experiments, primer extension products obtained by reverse transcription of purified rabbit liver tRNA^{Val} were subjected to nucleotide sequencing (not shown). These cDNAs showed exactly the sequences expected, i. e., they were complementary to nucleotides 1 to 5 of tRNA^{Val} with Val-6/20 as a primer and to nucleotides 1 to 21 with Val-22/41 as a primer (apart from the 15 and 20 nucleotides, respectively, of the primers themselves). In the latter case, modified nucleotides D and ψ were transcribed into A. These results indicated that faithfully transcribed cDNAs were to be expected, too, with pre-tRNAs as templates.

## pre-tRNA patterns in HeLa cells and human placenta

The patterns of primer elongation products characteristic for the two RNA sources are compared in Fig. 2. Some bands that are obtained with placenta RNA are absent in the case of HeLa cell RNA and *vice versa*. Other bands are obtained with both RNAs but vary in their intensity. Obviously, there are different pre-tRNA populations present in the two RNA preparations examined, indicating that several members of the tRNA^{Val} gene family exhibit different transcriptional activities in HeLa cells and placenta tissue.

Genes for the known tRNA^{Val} isoacceptors that are expressed *in vivo*

Fig. 3 shows a selection of autoradiograms obtained from the primer extension experiments described above. They display the nucleotide sequences of the 3' ends of the cDNAs which enabled us to deduce the 5' leader sequences of precursor tRNAs occurring *in vivo* that had been the templates for reverse transcription. Some of the products identified could be assigned unequivocally to tRNA^{Val} genes that have already been cloned and characterized by us (1 – 3) as well as by Craig *et al.* (5) and by Shortridge *et al.* (4). These genes are also indicated in Fig. 2.

cDNAs with the nucleotide sequence 5'...GAAACCAAC³' were found in HeLa cells and in placenta, and a cDNA 5'...GAAACCAACGGT³' was found in HeLa cells. The corresponding pre-tRNAs have the 5' terminal sequences 5'GUUGGUUUC...³' and 5'ACCGUUGGUUUC...³' and hence derive from the cellular tRNA^{Val} gene corresponding to that in clone pHtV1. A cDNA with the nucleotide sequence 5'...GAAACAGAT³' was detected in HeLa cells and in placenta. It indicates the presence of a precursor with the 5' terminal sequence 5'AUCUGUUUC...³' that is a transcript of the cellular counterpart of the tRNA^{Val} gene in clone pHtV8 (3) which is identical with the tRNA^{Val}$_{IAC}$ gene in clone λ hVKV7 (5). The genes in pHtV1 and pHtV8 both code for the major tRNA^{Val}$_{IAC}$ species.

A cDNA with the sequence 5'...GAAACCGGC³' in HeLa cells and a cDNA with the sequence 5'...GAAACCACCTT³' in
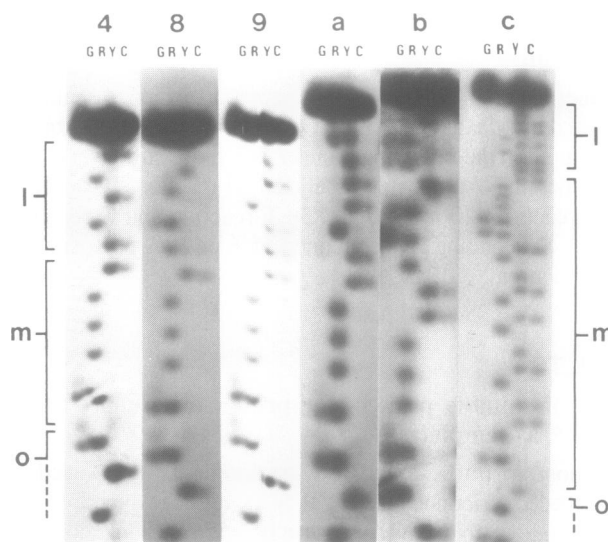
Figure 3: Sequence analysis of cDNAs obtained by primer extension of pre-tRNAs^{Val}. All major cDNA bands from the preparative polyacrylamide gel shown in Fig. 2 were sequenced according to the method of Maxam and Gilbert (15), and a selection of these analyses is shown. Nucleotides of the oligonucleotide primer (o), nucleotides complementary to the 5' terminal sequence of the mature tRNA^{Val} (m) and nucleotides complementary to the individual 5' leader sequences (9) can be read in 5' – 3' direction from bottom to top. Brackets on the left of the Figure refer to panels 4, 8, 9, a, and b, brackets on the right of the Figure refer to panel c. G: G-reaction; R: G+A-reaction; Y: C+T-reaction; C: C-reaction. Panels 4, 8, and 9: sequence analyses of cDNAs derived from pre-tRNAs of pHtV4-, pHtV8-, and pHtV9-related genes, respectively. (a): cDNAs from the pre-tRNA of the tRNA^{Val}$_{CAC}$ gene corresponding to that in clone λ hVKV7 (5). (b): cDNA from the pre-tRNA of a yet unknown tRNA^{Val} gene. The results for the clone λ hPTV5-related gene (panel c) were obtained in an experiment with the primer Val-22/41 (not shown). The results of all sequence analyses are summarized in Table 1.

HeLa cells and in placenta indicate the existence of pre-tRNAs with the sequences 5'GCCGGUUUC...3' and 5'AAGGUGGUUUC...3', respectively. These precursors should result from transcription of cellular genes related to the tRNA^Val_CAC genes of clones pHtV3 (1) and λ hVKV7 (5).

So, at least four genes for the two known tRNA^Val species are active *in vivo*. This appears reasonable, since for placenta it is known that the tRNA^Val population consists of the major isoacceptor species with an IAC anticodon and the minor species with a CAC anticodon in a 2:1 ratio (7). An activity *in vivo* of the pHtV6-analogous gene which also codes for tRNA^Val_IAC could not be demonstrated.

Variant tRNA^Val genes that are expressed *in vivo*

In addition to the four 'conventional' tRNA^Val genes, two human tRNA^Val allogenes were shown to be active in placenta. As deduced from the nucleotide sequence of the corresponding cDNA there is a precursor species in placenta RNA that has the 5' leader 5'GUCAGG...3' along with a 5'...GCUUC...3' motif corresponding to positions 1 to 5 of the mature tRNA. This precursor can only derive from a cellular allogene corresponding to that in pHtV5.

The expression of the second tRNA^Val allogene was detected by preparative primer extension with the oligonucleotide Val-22/41 as a primer (not shown; experimental details in Materials and Methods). Sequencing of one of the cDNAs synthesized revealed the presence of a pre-tRNA with the 5' terminal sequence 5'AGGUGUGGGUUUC...3'. Furthermore, this precursor has a C at position 20 of the mature tRNA, where conventional tRNAs^Val have a D. It obviously derives from the variant tRNA^Val gene related to that of clone λ hPTV5. An *in vivo* expression of the third tRNA^Val allogene which is represented by clone pHtV7 was not detected.

Pseudogenes expressed *in vivo*

Kahnt *et al.* (2) and Thomann *et al.* (3) have described three human tRNA^Val pseudogenes. As indicated by the corresponding cDNAs, there are RNA species with the 5' termini 5'AGGGUUGUUUC...3', 5'GUGUGGUUUC...3', and 5'AGUGGGUUUC...3' in unfractionated tRNA preparations from Hela cells or placenta. These are leader sequences that are expected for transcription products of the three pseudogenes represented by clones pHtV2, pHtV4, and pHtV9, respectively. This means that genes corresponding to any of the tRNA^Val pseudogenes characterized so far are transcribed *in vivo*, namely pHtV2- and pHtV4-related genes in HeLa cells as well as pHtV4- and pHtV9-related genes in placenta. The pre-tRNAs of the cellular genes corresponding to pHtV4 and pHtV9 are predominant species in the respective cell types studied.

Table 1 summarizes the available data concerning the expression of tRNA^Val genes in the two cell types examined and gives an evaluation for the abundance of the corresponding pre-tRNAs. Among the active genes there are representatives of any of the three classes of genes described by Thomann *et al.* (3):

a) genes for the two known tRNA^Val isoacceptors identified in human placenta (7);

b) gene variants, also termed 'allogenes' according to Leung *et al.* (8), which exhibit sequence alterations as compared to conventional genes, but which may nevertheless code for yet unknown functional, minor tRNA isoacceptors;

c) pseudogenes, which produce rather stable pre-tRNAs and apparently do not give rise to mature tRNAs (3).

Yet unknown tRNA^Val genes that are transcribed *in vivo*

Apart from the pre-tRNAs transcribed from already characterized tRNA^Val genes, some precursors occur in HeLa cell and placenta tRNA preparations that obviously derive from so far unknown tRNA^Val genes. In one case, the 5' leader sequence could be determined: it is 5'ACCGCCUG...3' (Fig. 3). However, in spite of our rather rigorous electrophoresis conditions there still existed the problem that certain cDNAs could not be sufficiently separated on the gel and therefore contaminated each other. Hence, the 5' leader sequences of the corresponding pre-tRNAs could not be determined unequivocally. Nevertheless, since they all contain a 5'...GUUUC...3' motif which was clearly distinguishable, they certainly are pre-tRNAs^Val.

The estimated number of the so far unknown tRNA^Val genes transcribed *in vivo* may be up to four. Therefore, together with the members of the tRNA^Val gene family active *in vivo* already mentioned above, whose number amounts to nine, up to thirteen tRNA^Val genes, allogenes, and pseudogenes appear to be transcribed *in vivo*. This means that, together with those genes whose expression *in vivo* could not be demonstrated, namely the pHtV6- and pHtV7-related genes, the tRNA^Val gene family contains at least fifteen independent members.

Traces of other pre-tRNAs were also detected by the primer extension experiments described. As judged by the sequences

Table 1. *In vivo* expression of human tRNA^Val genes.

Clone name (Refs.)	Sequences	H	P
	5' 3'		
pHtV1 (1)	...GTAACCGTTG gtttc...	+	+
pHtV2 (1)	...CAGCAGGGTT gtttc...	+	-
pHtV3 (1)	...CATCAAGCCG gtttc...	+	-
pHtV4 (2)	...TCCGCGTGTG gtttc...	++	++
pHtV5 (3)	...AAATGTCAGG gcttc...	-	+
pHtV6 (3)	...ACGCGTGCAA gtttc...	-	-
pHtV7 (3)	...ACGCCTACCG gtttc...	-	-
pHtV8 (3)	...CTCCGCATCT gtttc...	++	++
pHtV9 (3)	...CCCTTAGTGG gtttc...	-	++
Lambda hPTV5(tRNA^Val)(4)	...GTAGGTGTGG gtttc...	+	+
Lambda hVKV7(tRNA^Val_CAC)(5)	...AGCCAAGGTG gtttc...	+	+
unknown tRNA^Val gene	...ACCGCCUG gtttc...	+	+

Clone names and corresponding references of all sequenced tRNA^Val genes are given on the left. Flanking sequences of the non-coding, tRNA-like strands are shown in capital letters, 5' sequences of the structural genes in small letters. Arrowheads identify the transcription initiation sites of pre-tRNAs as determined by isolation and sequencing of the respective cDNAs in those cases where expression *in vivo* has been observed. As for the genes corresponding to those in clones pHtV1, pHtV4, and perhaps also in λ hVKV7 (arrowhead in parentheses), more than one initiation site has been detected. H = pre-tRNAs found in HeLa cells; P = pre-tRNAs found in placenta. ++ = 30 000 to 60 000 cpm, + = 4000 to 30 000 cpm per cDNA band; − = not found in the respective cell type. The sequence of the 5' flanking region and of the 5' terminus of a putative, yet unknown tRNA^Val gene is presented at the bottom.

corresponding to their mature 5' ends they were not transcribed from tRNA^{Val} genes. However, they only represent a small minority of the products and, furthermore, the corresponding cDNAs show only little radioactive incorporation. This demonstrates the specificity of the method of primer extension, leading to inefficient elongation of the tRNA^{Val}-derived primer with templates other than pre-tRNAs^{Val} or tRNA^{Val}.

Transcription initiation sites *in vivo*

In vitro, any of the nine human tRNA^{Val} genes examined by Thomann *et al.* (3) had more than one site for the initiation of transcription. As we show here, the initiation sites *in vivo* correlate with those found *in vitro* following a simple rule: The initiation site which is the most intensive *in vitro* is found in the precursor occurring *in vivo* (Table 1). Interestingly, in HeLa cells there are obviously three pre-tRNAs that may derive from the pHtV4-related cellular gene. They have the 5' terminal sequences 5'GCGUGUGGUUUC...3', 5'GUGUGGUUUC...3', and 5'GUGGUUUC...3'. Therefore, in the case of this gene initiation *in vivo* appears to occur at positions −7, −5, and −3 in exact agreement with the situation *in vitro* where initiation was also demonstrated to occur at these three sites (3). Initiation at position −5 is most frequent both *in vitro* and *in vivo*. Also, the gene corresponding to pHtV1 has two initiation sites in HeLa cells: −4 and −7 (see above). Finally, the precursor with the 5' terminus 5'GUGGUUUC...3' may also derive from the tRNA^{Val}$_{CAC}$ gene represented by clone λ hVKV7 (Table 1) which would mean that this gene, too, possibly has more than one initiation site *in vivo*. These results indicate that the finding of multiple initiation sites is not an *in vitro* artifact. A certain flexibility in the choice of the transcription initiation site seems to be a characteristic feature of RNA polymerase III.

DISCUSSION

So far, there is only limited information about the expression of tRNA gene families of higher eukaryotes *in vivo*. A major obstacle for studying this subject is the characteristic redundancy of genetic information for a given tRNA species. The expression of a single gene or of several members out of a group of genes coding for the same isoacceptor therefore can not easily be monitored, and there are only a few examples where this was indeed possible. Cytogenetic studies of chromosomal deletion mutants carried out at salivary gland polytene chromosomes provided evidence for the expression of tRNA^{Val} gene clusters in *D. melanogaster* (17). In yeast (18) and in *C. elegans* (19) the *in vivo* expression of tRNA^{Tyr} and tRNA^{Trp} genes, respectively, has been studied via natural suppressor mutants. Neither of these methods, however, can be used in the case of vertebrates. Here, a possible strategy is to demonstrate the activity of a tRNA gene by identifying its primary transcripts in the RNA population of a certain cell type or tissue.

Under these aspects it is a breakthrough that primer extension has been proven to detect individual pre-tRNA species and therefore can reveal the activity of single tRNA gene loci in cultured cells or tissues. This had not necessarily to be expected, since there are conflicting data concerning the feasibility of primer extension with tRNA templates. The method was successfully used by Dingermann and Nerke (20) to investigate tRNA gene expression in yeast, whereas another group reported difficulties especially with modified nucleotides as m^{1}A and wyeosine (21). Recently primer extension has been applied to identify a

developmentally regulated tRNA^{Val} gene in *Dictyostelium discoideum* (22) and to determine the steady-state levels of unspliced precursors produced by oocyte and somatic tRNA^{Tyr} genes in *X. laevis* (23).

We found that reverse transcription of human tRNA^{Val} and its precursors indeed yields faithfully transcribed cDNA products. It is certainly important that there are no highly modified nucleotides in that part of the pre-tRNAs^{Val} and tRNAs^{Val} which functions as the template for primer extension. In any such case this technique is an accurate, reproducible, and extremely sensitive method to study tRNA gene expression *in vivo*, detecting minimal amounts of precursor against a high, i. e. 500- to several 1000-fold background of mature tRNA.

In the course of our experiments we were able to demonstrate the activity of several members of the human tRNA^{Val} gene family in Hela cells and human placenta. It was expected that among them there are genes for the known tRNA^{Val} isoacceptor species isolated from human placenta (9). Indeed, two genes for the major tRNA^{Val}$_{IAC}$ are expressed: the cellular genes corresponding to those in pHtV1 and pHtV8 (1, 3), the latter of which is identical with the tRNA^{Val}$_{IAC}$ gene of clone λ hVKV7 (5). Also, two genes for the minor tRNA^{Val}$_{CAC}$ are transcribed, i. e. the cellular counterparts of the tRNA^{Val}$_{CAC}$ genes in clones pHtV3 (1) and λ hVKV7 (5). An activity of the gene corresponding to clone pHtV6 (3) which also codes for tRNA^{Val}$_{IAC}$ could not be detected. There are several possible reasons for this fact:

a) The gene is active, but its precursors belong to those whose corresponding cDNAs could not sufficiently be purified to yield unambiguous data upon nucleotide sequencing. In this case, improved purification methods should reveal an expression of these genes *in vivo*.

b) The gene is silent and does not contribute to the pool of mature tRNAs^{Val}, at least not in HeLa cells or in placenta.

c) The gene is active, but the concentration of the corresponding precursors is still too low to be detected even by the sensitive method of primer extension.

In principle, there are two processes that control pre-tRNA steady state concentrations in cells: the transcriptional activity of tRNA genes and the velocity of pre-tRNA maturation. pHtV6 is a clone with average transcriptional activity, but its primary transcripts are very rapidly processed *in vitro* (3). If their maturation occurs as rapidly *in vivo*, this may be the reason for a low precursor level.

So far there is no evidence that *in vitro* assays faithfully reproduce the expressional properties of tRNA genes *in vivo*. Recent results rather indicate that tRNA gene activity in organisms follows more complex rules than data obtained by *in vitro* experiments suggest. For example, the tRNA^{Tyr} genes of *C. elegans* are obviously controlled in a developmental stage- or tissue-specific manner (19). It is well established that tRNA pools adapt to protein synthesis requirements in highly differentiated cells (24, 25). Regulation phenomena like those observed in the nematode may be functional in other organisms, too, and may be involved in that adaptation by keeping an appropriate number of genes active to maintain exactly the concentration of a tRNA needed under certain circumstances. Therefore, tissue-specific factors may also be responsible for a low expression level or even inactivity of the pHtV6-related and other silent genes (see below) in placenta and may activate them in other tissues.

Those tRNAs for which a tissue- or stage-specific occurrence

has clearly been shown are all variant isoacceptors as, e. g., the silkgland-specific tRNA^{Ala} and tRNA^{Ser} of *B. mori* (26, 27), the bovine lens-specific tRNA^{Phe} (28), and the stage-specific tRNA^{Tyr} genes of *X. laevis* (23). The most extensively studied example is that of the *B. mori* tRNA^{Ala}; here, a sequence element in the 5' flanking region of the corresponding variant gene is responsible for an expression restricted to the silkgland (11). For *D. melanogaster*, variant tRNA genes have also been well characterized and termed 'allogenes'. Although they are often more efficiently transcribed *in vitro* than conventional genes (8), an activity *in vivo* has not been detected so far (17). These data suggest that variant tRNA genes in general might be candidates for tissue-specifically regulated genes.

Although the two known valine-accepting tRNAs already mentioned suffice to read all four valine codons, at least one more tRNA^{Val} species obviously exists in human placenta (9) which has not yet been purified and sequenced. There is the possibility that this species differs from the other two only by the degree of nucleotide modification. On the other hand it is not unlikely that one of the human tRNA^{Val} allogenes may code for it. Indeed, tRNA^{Val} genes corresponding to those in clones pHtV5 and λ hPTV5 were shown to be expressed in placenta. Which of them gives rise to the third tRNA^{Val} can only be elucidated by isolation and sequencing of the corresponding isoacceptor.

Since, as already mentioned the known allogenes of *D. melanogaster* could not be shown to be active *in vivo* (17) and since variant tRNAs may be restricted to highly specialized tissues (26–28) or certain developmental stages (22,23), this readily detectable activity of two tRNA^{Val} allogenes is a surprise. It may mean either that allogenes of the human tRNA^{Val} gene family can be constitutively expressed or that placenta is one of those specialized tissues that require the expression of variant tRNA genes.

As for the pHtV7-related, third tRNA^{Val} allogene, whose activity *in vivo* we also wanted to show, a reverse transcript derived from a corresponding pre-tRNA was not found among the sequenced cDNAs. Since the precursor could have escaped detection for one of the reasons discussed above, we intended to demonstrate the presence of the variant tRNA itself because of its presumably higher concentration as compared to that of its precursor. We utilized an oligonucleotide primer for reverse transcription that has 100% complementarity with the possible pHtV7-related tRNA^{Val} but which exhibits two mismatches causing inefficient priming when hybridizing to the conventional tRNAs^{Val}, since the region of homology spans that part of the T arm where the tRNA of this allogene deviates from the known tRNAs^{Val} (3). In several experiments we could not detect the occurrence of the respective variant tRNA in placenta but we could not completely rule out its presence in HeLa cells. These results may mean that the cellular counterpart of the pHtV7 gene is not principally inactive but that it is possibly a candidate for a differentially regulated tRNA gene.

Our results demonstrate with remarkable clarity different pre-tRNA^{Val} populations in HeLa cells and placenta. It appears unlikely that these differences are caused by different levels of processing enzymes, since certain pre-tRNAs are absent or low in one of both cell type. Our data indicate that primer extension may be a powerful general approach for studying stage- or tissue-specific tRNA gene expression. Detailed statements about differential tRNA gene expression as presented here are preferentially feasible in those cases, where the genes and the respective flanking sequences for most or all members of a tRNA

isoacceptor family are known and characterized by *in vitro* studies, and if the 5' leaders of individual pre-tRNAs are so different that the sequences between the mature 5' end and the transcription start points are unique.

Kahnt *et al.* (2) and Thomann *et al.* (3) have described three tRNA^{Val} pseudogenes, those in clones pHtV2, pHtV4, and pHtV9. Since neither of them has defects in the internal promoter boxes, they are transcribed in cell-free transcription systems, however, they do not produce mature tRNAs. The primary transcripts of pHtV4 and pHtV9 are not processed at all; pHtV2-derived transcripts are correctly processed at their 5' ends whereas at the 3' end the precursor is cut 3 or 4 nucleotides downstream from the mature tRNA domain and, consequently, the CCA terminus is not added to this immature tRNA.

Considering the transcriptional activity of the pseudogenes *in vitro*, it was particularly interesting to find out whether they are also active *in vivo*. Indeed, this is the case for each of the three pseudogenes. It is one of the unexpected surprises of this study that their transcripts appear to accumulate and to be prominent species within the pre-tRNA population. This provokes the question what the significance of these transcripts may be *in vivo*. Are they produced in order to serve some yet unknown function as suggested, e. g., by Zasloff *et al.* (29)? According to their model, unprocessed or slowly processed tRNA precursors which can not enter the cytoplasm might serve as a feedback signal to sense the transcriptional output of tRNA gene families. Only further experiments can bring an answer to these and other open questions concerning tRNA gene families and their expression.

ACKNOWLEDGEMENTS

REFERENCES

1. Arnold, G. J., Schmutzler, C., Thomann, U., van Tol, H., and Gross, H. J. (1986) *Gene* 44, 287–297.
2. Kahnt, B., Frank, R., Blöcker, H., and Gross, H. J. (1989) *DNA* 8, 51–58.
3. Thomann, H.-U., Schmutzler, C., Hüdepohl, U., Blow, M., and Gross, H. J. (1989) *J. Mol. Biol.* 209, 505–523.
4. Shortridge, R. D., Johnson, G. D., Craig, L. C., Pirtle, I. L., and Pirtle, R. M. (1989) *Gene* 79, 309–324.
5. Craig, L. C., Wang, L. P., Lee, M. M., Pirtle, I. L., and Pirtle, R. M. (1989) *DNA* 8, 457–471.
6. Hatlen, L. and Attardi, G.(1971) *J. Mol. Biol.* 56, 535–553.
7. Chen, E. Y. and Roe, B. A. (1977) *Biochem. Biophys. Res. Commun.* 78, 631–640.
8. Leung, J., Addison, W. R., Delaney, A. D., Mackay, R. M., Miller Jr., R. C., Spiegelman, G. B., Grigliatti, T. A., and Tener, G. M., (1984) *Gene* 34, 207–217.
9. Anandaraj, M. P. J. S. and Roe, B. A. (1975) *Biochemistry* 14, 5068–5073.
10. Zasloff, M., Santos, T., and Hamer, H. (1982) *Nature* 295, 533–535.
11. Young, L. S., Takahashi, N., and Sprague, K. U. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 374–378.
12. Jank, P., Shindo-Okada, N., Nishimura, S., and Gross, H. J. (1977) *Nucleic Acids Res.* 4, 1999–2008.
13. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
14. Roe, B. A. (1975) *Nucleic Acids Res.* 2, 21–42.
15. Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560.

16. Sprinzl, M., Hartmann, T., Weber, J., Blank, J., and Zeidler, J. (1989) *Nucleic Acids Res.* 17, r1−r172.
17. Larsen, T. M., Miller Jr., R. C., Spiegelman, G. B., Hayashi, S., Tener, G. M., Sinclair, D. A. R., and Grigliatti, T. A. (1982) *Mol. Gen. Genet.* 185, 390−396.
18. Olson, M., Loughney, V. K., and Hall, B. D. (1979) *J. Mol. Biol.* 132, 387−410.
19. Kondo, K., Hodgkin, J., and Waterston, R. H. (1988) *Mol. Cell. Biol.* 8, 3627−3635.
20. Dingermann, T. and Nerke, K. (1987) *Analyt. Biochem.* 162, 466−475.
21. Wittig, B. and Wittig, S. (1978) *Nucleic Acids Res.* 5, 1165−1178.
22. Dingermann, T., Amon-Böhm, E., Bertling, W., Marschalek, R., and Nerke, K. (1988) *Gene* 73, 373−384.
23. Stutz, F., Gouilloud, E., and Clarkson, S. G. (1989) *Genes Dev.* 3, 1190−1198.
24. Garel, J. P., Mandel, P., Chavancy, G., and Dailli, J. (1970) *FEBS Lett.* 7, 327−329.
25. Hatfield, D., Varricchio, F., Rice, M., and Forget, B. G. (1982) *J. Biol. Chem.* 257, 3183−3188.
26. Hentzen, D., Chevallier, A., and Garel, J. P. (1981) *Nature* 290, 267−269.
27. Sprague, K. U., Hagenbüchle, O., and Zuniga, M. C. (1977) *Cell* 11, 561−570.
28. Lin, F. K., Furr, T. D., and Chang, S. H. (1980) *J. Biol. Chem.* 255, 6020−6024.
29. Zasloff, M., Rosenberg, M., and Santos, T. (1982) *Nature* 300, 81−84.